Mapping between prosodic hierarchy and supralaryngeal articulatory variations in French

Yohann Meynadier

Mapping Between Prosodic Hierarchy and Supralaryngeal Articulatory Variations in French

Yohann Meynadier

Laboratoire Parole et Langage
Université de Provence, Aix-en-Provence, France
yohann.meynadier@lpl.univ-aix.fr

Abstract

The gradual behaviour of articulatory supralaryangeal variations as a function as an 8-level *vs* a 4-level prosodic hierarchy was analysed here. Comparisons between the results related to each hierarchy suggested that 4-level hierarchy was sufficient to account for prosodic-dependent articulatory changes in French. These results allowed to discuss the architecture of hierarchical prosodic representation of speech.

1. Introduction

In this study [1], an investigation was conducted on the degree of mapping between articulatory supralaryngeal variations at prosodic boundaries and prosodic hierarchies, according to the nature and the number of levels.

Many works have established various kinds of articulatory strengthening of segments and of modifications of the dynamics of the intra-/inter-segmental articulation depending on the levels of boundary according to prosodic hierarchies. They yielded that gesture and coordination of segments vary in multi-dimensional ways, but only a one-way direction of segmental strengthening and reduction of coarticulation in initial, final or cross-boundary position of prosodic constituent of increasing hierarchical level. So, prosodic-dependant supralaryngeal correlates are able to distinguish up to 5 hierarchical levels of prosodic constituency [2; 3; 4; 5; 6].

Table 1: *Examples of the 8 prosodic boundaries. (ICC contexts were uttered in interrogative modality, because speakers were unable to produce statements without pause)*

LPH	*NPH*	Sentences
UNA	US	Ma belle-sœur a gagné deux qu**a**tre-**qua**tre noirs et une moto. *(My sister-in-law won two four-wheel drive and a moto.)*
	UW	Ma grand-mère utilise un s**ac** **Ta**ti pour faire ses courses. *(Ma grand-mother uses a Tati bag for shopping.)*
ACC	AW	La fatigue affaiblit l'att**aque** **la**otienne de l'équipe. *(Fatigue weakens the Laotian attack of the team.)*
	AP	La merveilleuse nounou ét**a**l**e** **Ca**mélia sur le lit. *(The wonderful nurse lays out Camelia on bed.)*
ICT	It	La réponse de la candid**a**t**e** **ca**tastrophe le jury. *(The answer of the candidate is stunning the jury.)*
	IT	Ma mauvaise foi et mes att**aqu**es, **Ta**tiana les méprise. *(My dishonesty and my attacks, Tatiana scorns them.)*
ICC	Ic	A force, elle les adore, les d**att**es, **Ca**mélia, maintenant ? *(Will she end up loving them, the dates, Camelia, now ?*
	IC	Ma belle-mère a trouvé l'ét**a**l ? **Ca**mélia le cherchait. *(Did my mother-in-law find the stall? Camelia looked for it.)*

The present study was not focused on the articulatory nature of segmental correlates of the prosodic hierarchy, but related to the granularity of these correlates depending on two prosodic hierarchies: (1) a *Narrow Prosodic Hierarchy* composed of numerous and fine/narrow hierarchical prosodic levels, that is detailed prosodic categories; (2) a *Large Prosodic Hierarchy* defined by more limited and rougher/larger hierarchical prosodic levels, that is a basic prosodic constituency. Statistical comparisons between these two kinds of prosodic hierarchies and the prosodic levels significantly distinguished by gradual articulatory marks of boundaries were made. The comparisons allowed to assess the maximal correlation degree between prosodic-dependant articulatory variations and the richness of prosodic hierarchy. The goal was to study if a rich hierarchical structure of speech is needed to account for phenomena of prosodic-dependant articulatory variation. In other words, looking at the prosodic-dependant articulatory variations could be a good mean to assess the depth of architecture of prosodic representations, viewed as a structure hierarchically organised on *n* levels of parallel segmentation of speech ([7], for a review).

2. Method

2.1. Speech material

Prosodic-dependant articulatory variations were observed in aC#Ca sequences embedded in a sentence, where CC stands for /kl/, /lk/, /kt/ or /tk/ and where # symbolises a prosodic boundary of different hierarchical levels. Accent always fell on the first vowel /a/ of the sequence, which is in the final syllable of prosodic constituent.

7 sentences for /kl/ and 8 for /lk, kt, tk/, each with a different inter-consonantal prosodic boundary, were read without pause 12 or 15 times by 3 French speakers (Table 1).

2.2. Prosodic hierarchies

2.2.1. Levels of prosodic hierarchies

Two kinds of hierarchy of prosodic categories were studied: *Narrow Prosodic Hierarchy (NPH)* vs *Large Prosodic Hierarchy (LPH)* (see Table 1 for an illustration). These different prosodic contexts were obtained by manipulating syntactic and thematic structure of sentences.

NPH was composed of 8 levels of narrow prosodic category, from the lower to the higher: (1) *unaccented syllable (US)*; (2) *unaccented word (UW)*; (3) *accented word (AW)*, i.e. a rhythmical final accent of word in medial position of an accentual phrase; (4) *accentual phrase (AP)*, i.e. a final accent of prosodic word; (5) *continuation intonational phrase (It)*, i.e. a final intonation of the subject phrase; (6) *high continuation intonational phrase (IT)*, i.e. a final intonation of a preposed thematic constituent; (7) *minor conclusion intonational phrase (Ic)*, i.e. a final intonation of a postposed thematic constituent; (8) *full/major conclusion intonational phrase (IC)*, i.e. a final intonation of an utterance.

NPH (*US* < *UW* < *AW* < *AP* < *It* < *IT* < *Ic* < *IC*) mainly came from the morphological model of French prosody [8]. Prosodic categories were considered as accentual or

intonational morphemes defined by a function, i.e. lexical, rhythmical/phonological, syntactic or pragmatic, and a form, i.e. suprasegmantal parameters: duration, intensity and F0 contour. Both pragmatic and syntactic hierarchies mainly determined the hierarchical prosodic structure of utterances. Hence, this conception was closer to the morphosyntactic phonological approaches of hierarchical prosodic constituency [9; 10] than pure intonational models.

LPH grouped two by two the narrow prosodic categories of *NPH*. *LPH* was composed of 4 levels of large prosodic category, from the lower to the higher: (1) *syllable/word* (*UNA*), i.e. *US+UW* from *NPH*; (2) *accentual group (ACC)*, i.e. *AW +AP*; (3) *continuation intonational group (ICT)*, i.e. *It+IT*; (4) *conclusion intonational group (ICC)*, i.e. *Ic+IC*.

LPH (*UNA* < *ACC* < *ICT* < *ICC*) was closer to pure prosodic hierarchies proposed by intonational models of prosodic structure for French as [11] or [12]. Nevertheless, *LPH* differed from them in several respects. In these models, the lexical unit is not considered as a prosodic unit, but the word can be seen as the minimal accentuable unit in French. A final rhythmical accent (here, *AW*) is a prosodic mark of the *Tonal Unit* [11] or is a secondary prominence of the *Accentual Phrase* [12]. So, the large accentual category, i.e. *ACC* was heterogeneous regarding the demarcative function of its narrow accentual categories: *AW* had none as opposed to *AP*. Here, only the accentual (*vs* unaccentued or *vs* intonational) nature of these prominences were the main base of this large prosodic level. Finally, continuation *vs* conclusion intonation were considered having no hierarchical relationships in these models. However, an intermediate unit between the accentual and the (full) intonational units has already been suggested. For example, [12] postulated that utterances with a preposed or postposed thematic constituent could be represented as two *Intermediate Intonation Phrases* ("ip") within an *Intonational Phrase*. [11] proposed that recursivity at intonation level could account for a left dislocated constituent as a *segment of Intonation Unit* of a bigger *Intonation Unit*. I postulate here that this levelling between intermediate *vs* full intonation groups could partly cover a functional distinction between continuation *vs* conclusion intonation. Moreover, this hierarchical relationship is taken in part in a morphological approach [8]. And, if the utterance can be seen as the largest prosodic constituent [9], only a conclusion intonation can mark its final boundary.

2.2.2. Suprasegmental evidences of prosodic hierarchies

Suprasegmantal realisation of these two prosodic hierarchies was assessed by means of three temporal and tonal measurements made on the first vowel /a/ of the sequence: (1) *vowel lengthening* as the lengthening ratio (in %) in relation to the average duration of unaccented /a/ (anywhere except in the first position of the sequence); (2) *F0 peak* as the maximal F0 value during the first /a/ of the sequence (in ERB normalised by *z* transform); (3) *vowel glissando* as the ratio between the pitch increase during the vowel (in ERB normalised by *z* transform) and the vowel duration.

Statistical results (Fig. 1) showed that both prosodic hierarchies were largely legitimated by these suprasegmental parameters, and particularly by vowel glissando which combined temporal and tonal information.

2.3. Articulatory measurements

An articulatory analysis of supralaryngeal correlates of prosodic hierarchies was based on an electropalatographic (EPG) investigation in temporal and spatial dimensions of linguopalatal gestures in aC#Ca sequences. 178 EPG measurements are done for each item:

- intra-gesture measurements on vowels (i.e. duration and amplitude of maximal opening; vowel gesture duration), and on consonants (i.e. absolute and relative duration of gesture, closing, closure, maximal constriction and opening phases; temporal symmetry of consonant; amplitude of maximal constriction; anteriority / posteriority and centrality of constriction; etc.)
- inter-gesture measurements on aC#, a(C)#C, #Ca, and C#(C)a coarticulation (i.e. absolute and relative duration of time interval between C gesture, closure or maximal constriction and V gesture or maximal opening; maximal difference of linguopalatal contact between C and V; etc.); and on C#C coarticulation (i.e. temporal proportion of C in C#C; absolute and relative duration of time interval between C1 and C2 gestures, closures or maximal constrictions; magnitude of spatial and spatio-temporal overlaps of C gestures; maximal spatial distance between places of articulation; etc.).

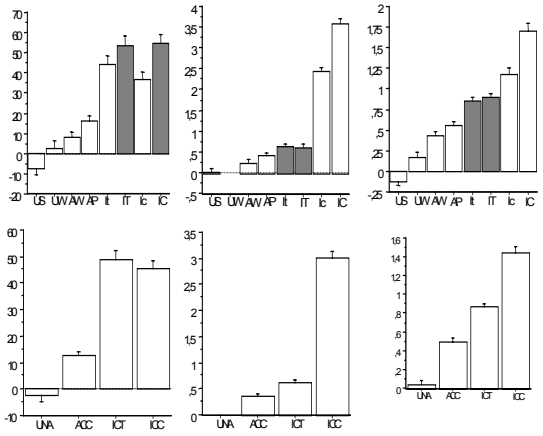

Figure 1: *Vowel lengthening (left), F0 peak (middle), and glissando (right) according to NPH (up) and LPH (down). In grey, non significant differences (PLSD Fisher post hoc tests with speaker data pooled).*

2.4. Statistical analysis

To assess if prosodic-dependant linguopalatal correlates are better accounted by a rich prosodic hierarchy or by a basic one, a multi-step selection procedure of articulatory results based on four statistical criteria was done according to both hierarchies.

In step #1, only articulatory measurements (n = 2136, i.e. 178 measurements * 4 CC clusters * 3 speakers) showing a significant effect of prosodic categories (i.e. one-factor ANOVA, $p < .05$, speakers split) were kept for the next step.

Step #2 consisted in excluding articulatory measurements not sufficiently correlated to prosodic hierarchies. Spearman rank correlation coefficient (r_s) gave a numeric distance of a one-to-one mapping between two alphanumeric variables by quantifying the relationship between these rankings. For each measurement, the averages of prosodic categories were ranked in increasing order (henceforth, *"articulatory hierarchy"*) which was compared, by means of r_s, with the prosodic hierarchy of reference, that is *NPH* or *LPH*. A

threshold of rejection was chosen to rule out the articulatory hierarchies too far from the expected prosodic ones. For *LPH*, it accounted for the impact of only two order inversions between two hierarchically contiguous prosodic categories ($-.6 < r_s < .6$). For *NPH*, the threshold was calculated to produce the same theoretical rejection level of articulatory hierarchies than for *LPH* comparisons ($-.36 < r_s < .36$).

Step #3 dealt with a comparison between speakers. To be selected, a measurement must be represented by at least two speakers out of three and, these individual articulatory hierarchies must all show the same ranking direction.

In final step #4, a one-factor ANOVA with data pooled over the remaining speakers was carried out with a rejection threshold at .05 significant level for each measurement. This was done to check the statistical homogeneity of speaker behaviours.

So, the goal of the selection procedure was to draw only the supralaryngeal correlates showing a close co-variation with the prosodic hierarchies and corresponding to a homogeneous inter-speaker behaviour.

3. Results

3.1. Quantitative comparison between the two hierarchies

NPH and *LPH* were compared according to the number of articulatory measurements selected at each step of the statistical procedure.

On the one hand, this provided information about the frequency of inter-speaker supralaryngeal articulatory correlates of the hierarchical prosodic structure of utterances. On the other hand, it informed on the possibility that a richer prosodic hierarchy is able to account for subtler prosodic-dependent articulatory variations than a crude one, as in the case of a largely greater number of articulatory phenomena selected with reference to *NPH* than to *LPH*.

Table 2: *Cumulative and relative* ratios of rejection (%) at each step of the selection procedure for NPH and LPH. * according to the number of remaining measurements from the previous step.*

steps of the procedure	cumulative ratio		relative ratio	
	NPH	*LPH*	*NPH*	*LPH*
step #1	29	37	29	37
step #2	57	64	40	43
step #3	77	81	40	30
step #4	78	81	3	.01

First, the results of Table 2 show that, at the end of the selection procedure, 22 % (i.e. 78 % rejected) and 19 % (i.e. 81% rejected) of the linguopalatal measurements were statistically close correlated with the prosodic hierarchies, *NPH* and *LPH* respectively. It suggests that these kinds of phonetic expression of prosodic structure were not anecdotal. And they were not due to chance factor. The relative ratio of rejection at step #2 was 15-18 % lower than the theoretical ratio due to random selection (58.3 %, cf. §2.4.). This confirms the point of view that the prosodic-dependant supralaryngeal phenomena embrace a general and inter-speaker behaviour in speech.

Moreover, there were no actual quantitative differences between the amount of different articulatory variations captured by *NPH* and by *LPH*. At the end of the selection procedure (step #4, even #3), only 3 % of articulatory measurements were captured by *NPH* and not by *LPH*. So, at this step of analysis, it can not be said that a richer prosodic hierarchy quantitavely better accounted for supralaryngeal variations than a poorer one. In other words, *LPH* seemed as efficient as *NPH* to capture this kind of prosodic-dependant correlates.

3.2. Qualitative comparison between the two hierarchies

To assess the qualitative contribution of *NPH* compared with *LPH* to describe prosodic-dependant articulatory variations, only measurements selected by both hierarchies were considered, i.e. 102 out of 157 for *NPH* and out of 133 for *LPH*. The comparison consisted in testing whether the finer hierarchy accounted for intra-level variations of the cruder one. In other words, whether *LPH* can be strongly sub-specified (i.e. detailed) by *NPH* concerning prosodic-dependant supralaryngeal correlates.

Pairwise comparisons between the *NPH* categories were made on the 102 articulatory measurements using PLSD Fisher post hoc tests pooling the remaining speakers. Sub-specifications of each large prosodic category of *LPH* by the two composing narrow categories of *NPH* were assessed by means of summing over the 102 measurements the number of cases where the two *NPH* categories were adjacent and significantly different ($p < .05$) or non adjacent and significantly different. The results are presented in Table 3.

Table 3: *Frequency (%) of cases of hierarchically adjacent and non adjacent NPH categories expressed in the remaining articulatory hierarchies*

% of cases	US-UW	AW-AP	It-IT	Ic-IC
Adjacent	56	53	41	45
and sign. different	*10*	*6*	*2*	*16*
Non adjacent	44	47	60	55
and sign. different	*28*	*33*	*32*	*30*

It appears that between 44 and 60 % of the pairs of *NPH* categories were not contiguous in the 102 articulatory hierarchies (3rd line of Table 3). Intonational narrow categories seemed less hierarchically linked than unaccented and accented ones. Only few (2 to 16 %) of them were significantly distinct prosodic categories. Moreover, about a third of non adjacent categories in *NPH* pairs were significantly distinguished. These facts revealed that it could not be said that *NPH* brought a strong profit for the account of fine details of prosodic-dependant articulatory correlates in comparison with *LPH*.

Two reasons can explain these results. First, *LPH* may be sufficient to account for the supralaryngeal variations; and so, *NPH* is not necessary because it provides a too detailed hierarchy causing much noise in the results. Secondly, the relationships between *NPH* and *LPH* levels may not match the correspondences suggested here.

In any case, in this study the large prosodic hierarchy seemed to give quantitatively and qualitatively to be a better account for prosodic-dependant supralaryngeal variations.

3.3. Prosodic granularity of supralaryngeal variations

To assess the validity of *LPH* to capture supralaryngeal variations at prosodic boundaries, a comparison was done with Fougeron's recent study on French [2]. She analysed variations of linguopalatal articulation (EPG) and of nasal

airflows of /t, k, s, l, n, i, ɑ̃/ segments in initial position of prosodic constituents of 4 or 5 hierarchical levels, from the lower to the higher: (1) *syllable (S)*; (2) *word (W)*; (3) *accentual phrase (AP)*; (4) *intonational phrase (IP)*; (5) *utterance (U)*.

This comparison between the two studies showed that the supralaryngeal articulatory correlates co-varied in a same way with both prosodic hierarchies (Table 4). In both studies, the prosodic-dependant supralaryngeal phenomena systematically reflected that the distinctions between the hierarchical levels obeyed to the same decreasing scale of prosodic boundary strength.

Table 4: *Frequency (%) of significant distinctions, of hierarchical inversions, without and with significant difference ($p < .05$), between prosodic levels each other here (M) and in Fougeron's work (F). * only inversions between AP and W.*

type of distinction		sign.distinction		inversion		sign.inversion	
		M	F	M	F	M	F
ICC/U	≠ ICT/IP	43	20	42	10	9	10
	≠ ACC/AP	*87*	*90*	-	-	-	-
	≠ UNA/SW	*95*	*100*	-	-	-	-
ICT/IP	≠ ACC/AP	77	68	3	2	0	2
	≠ UNA/SW	*95*	*86*	-	-	-	-
ACC/AP	≠ INA/SW	52	45	7	7*	1	7*

First, intonational units (i.e. both continuative and conclusive ones) showed a very strong hierarchical dominance over the other kinds of categories. In 77 to 100 % of the cases, the intonational levels were significantly marked by stronger articulatory correlates than the accentual and syllable/word levels. Moreover, hierarchical inversions between the intonational and non intonational levels were quasi nil (3 or 2 %).

Secondly, a less strong degree of hierarchical distinction concerned the dominance of accentual unit over syllable/word ones. 45 % of the cases in this study and 52 % in Fougeron's showed that the accentual level were produced with stronger articulatory variations than the syllable/word ones. Only few cases presented a hierarchical inversion between these levels.

Finally, with a little discrepancy of frequency between both studies, the lower hierarchical distinction was related to a dominance of the conclusive intonation boundary over the continuative one. In 20 and 40 % of the cases, the degree of supralaryngeal variations was more important for the conclusive level than the continuative one. Furthermore, regarding this distinction inside the intonational level, more frequent hierarchical inversions were observed for this distinction than the other kinds of hierarchical distinction.

4. Discussion and Conclusion

This work was an attempt to show how testing different mappings between different kinds of prosodic hierarchies according to supralaryngal articulatory variations could be a good way to assess, independently from the only suprasegmental features, the depth of hierarchical levels implied to prosodically structure utterances. The results obtained are largely shared by previous studies, particularly on French, concerning the granularity of prosodic hierarchy according to prosodic-dependent supralaryngeal variations.

On the one hand, the results of this study fit in with the general idea of a need for another hierarchical level in the representation of prosodic structure of utterances, either above or below intonational unit [7; 11; 12]. For instance, a hierarchical levelling between the conclusive and continuative intonation boundaries emerged, even if this hierarchical distinction was optional and relatively instable regarding the articulatory variations. Although not represented in pure intonational models, the conclusive *vs* continuative distinction may participate to determine an added hierarchical prosodic level.

On the other hand, this work argues for hierarchical prosodic representation with a limited number of levels, that is closer to purely intonational conceptions of prosodic hierarchy presenting less rich constituency [3; 11; 12] than the morphosyntactic models counting up to 7-8 levels [9; 10].

There are no general consensus about the nature, the number and the relationships of levels of prosodic hierarchy today [7; 12]. Taking into account the interplay between articulatory manoeuvres and prosodic structure can provide another mean to assess and discuss the details of architecture of prosodic representation of speech structure.

This point of view argues for more unified conceptions between the segmental and suprasegmental dimensions in phonological theories of speech production [3; 4].

5. References

[1] Meynadier, Y., 2003. *Interaction entre prosodie et (co)articulation linguopalatale en français*. Université de Provence: PhD dissertation.
< http://www.lpl.univ-aix.fr/lpl/documents/theses >

[2] Fougeron, C., 1998. *Variations articulatoires en début de constituants prosodiques de différents niveaux en français*. Université Paris III: PhD dissertation.

[3] Beckman, M.; Edwards, J.; Fletcher, J., 1992. Prosodic structure and tempo in sonority model articulatory dynamics. In *Papers in Laboratory Phonology I*, G. Docherty & R. Ladd (eds.) Cambridge: CUP, 68-86.

[4] Byrd, D., 2000. Articulatory vowel lengthening and coordination at phrasal junctures. *Phonetica* 57, 3-16.

[5] Cho, T., 2001. *Effects of Prosody on Articulation in English*. UCLA: PhD dissertation.

[6] Keating, P.; Cho, T., Fougeron, C.; Hsu, C.H., 2003. Domain-initial articulatory strengthening in four languages. In *Papers in Laboratory Phonology VI*, J. Local, R. Odgen & R. Temple (eds). Cambridge: CUP.

[7] Shattuck-Hufnagel, S.; Turk, A., 1996. A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* 25(2), 193-247.

[8] Rossi, M., 1999. *L'intonation, le système du français : description et modélisation*. Gap: Ophrys.

[9] Nespor, M.; Vogel, I., 1986. *Prosodic Phonology*. Dordrecht: Foris Publications.

[10] Selkirk, E., 1984. *Phonology and Syntax: the Relation between Sound and Structure*. Cambridge: MIT Press.

[11] Di Cristo, A., 1998. French intonation. In *Intonation Systems*, D. Hirst & A. Di Cristo (eds). Cambridge: CUP, 195-218.

[12] Jun, A-S.; Fougeron, C., 2000. A phonological model of French intonation. In *Intonation: Models and Technology*, A. Botinis (ed.). Dordrecht: Kluwer, 209-242.

Acknowledgements

This work was supported by a grant from Conseil Régional PACA. Many thanks to Michel Pitermann (LPL) for his help to statistical analysis.